

**Analisis Data Status Di Facebook Untuk
Merekomendasikan Produk Makanan Ringan Menggunakan
TF-IDF Dan *Document Weighting***

JURNAL

**Diajukan Kepada
Fakultas Teknologi Informasi UKSW
Untuk Memperoleh Gelar Magister Komputer**



**Oleh :
Rohmad Abidin
NIM : 972016100**

**Program Studi Magister Sistem Informasi
Fakultas Teknologi Informasi
Universitas Kristen Satya Wacana
Salatiga
Oktober 2018**



PERNYATAAN TIDAK PLAGIAT

Sebagai sivitas akademik Universitas Kristen Satya Wacana, saya yang bertanda di bawah ini :

Nama : Rohmad Abidin
NIM : 972016100 Email : rohmada@stekom.ac.id
Fakultas : Teknologi Informasi Program Studi : Magister Sistem Informasi

Judul Tugas Akhir : Analisis Data Status Di Facebook Untuk Merekomendasikan Produk Makanan Ringan Menggunakan *TF-IDF* Dan *Document Weighting*.

Pembimbing : 1. Prof. Dr. Ir. Eko Sedyono, M.Kom
2. Dr. Sri Yulianto Joko Prasetyo, S.Si., M.Kom

Dengan ini menyatakan bahwa :

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Kristen Satya Wacana maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Kristen Satya Wacana.



Salatiga, 16 Oktober 2018

Rohmad Abidin



PERPUSTAKAAN UNIVERSITAS
UNIVERSITAS KRISTEN SATYA WACANA

Jl. Diponegoro 52-60 Salatiga 50711

Jawa Tengah, Indonesia

Telp. 0298 – 321212, Fax 0298 321433

Email : library@adm.uksw.edu ; http://library.uksw.edu

PERNYATAAN PERSETUJUAN HAK AKSES

Sebagai sivitas akademik Universitas Kristen Satya Wacana, saya yang bertanda di bawah ini :

Nama : Rohmad Abidin
NIM : 972016100 Email : rohmad@stekom.ac.id
Fakultas : Teknologi Informasi Program Studi : Magister Sistem Informasi

Judul Tugas Akhir : Analisis Data Status Di Facebook Untuk Merekomendasikan Produk Makanan Ringan Menggunakan *TF-IDF* Dan *Document Weighting* .

Dengan ini saya menyerahkan hak *non-eksklusif** kepada Perpustakaan Universitas – Universitas Kristen Satya Wacana untuk menyimpan, mengatur akses serta melakukan pengolahan terhadap karya saya ini dengan mengacu pada ketentuan akses tugas akhir elektronik sebagai berikut (beri tanda pada kotak yang sesuai) :

- ☒ a. Saya mengizinkan karya saya tersebut diunggah ke dalam aplikasi koleksi Digital Perpustakaan Universitas dan/atau portal GARUDA.
- ☐ b. Saya tidak mengizinkan karya tersebut diunggah ke dalam aplikasi koleksi Digital Perpustakaan Universitas, dan/atau portal Garuda. **

* Hak yang tidak terbatas hanya bagi satu pihak saja, Pengajar, Peneliti dan mahasiswa yang menyerahkan hak non-eksklusif kepada Repositori Perpustakaan Universitas saat mengumpulkan hasil karya mereka masih memiliki hak copyright atas karya tersebut.

** Hanya akan menampilkan halaman judul dan abstrak. Pilihan ini harus dilampiri dengan penjelasan/alasan tertulis dari pembimbing TA dan diketahui pemimpin fakultas (dekan/kaprogdi).

Dengan pernyataan ini saya buat dengan sebenarnya.

Salatiga, 16 Oktober 2018

Rohmad Abidin

Mengetahui,

Prof. Dr. Ir. Eko Sedyono, M.Kom
Pembimbing I

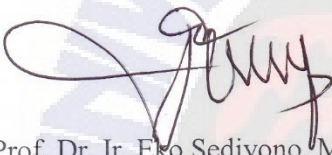
Dr. Sri Yulianto Joko Prasetyo, S.Si., M.Kom
Pembimbing II

Lembar Pengesahan


Judul Artikel : Analisis Data Status Di Facebook Untuk Merekomendasikan Produk Makanan Ringan Menggunakan *Tf-Idf* Dan *Document Weighting*.

Nama Mahasiswa : Rohmad Abidin
NIM : 972016100
Program Studi : Magister Sistem Informasi
Fakultas : Teknologi Informasi

Menyetujui,

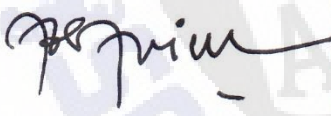


Prof. Dr. Ir. Eko Sedyono, M.Kom
Pembimbing I




Dr. Sri Yulianto Joko Prasetyo, S.Si., M.Kom
Pembimbing II

Mengesahkan,



Wiwin Sulistyo, ST., M.Kom
Dekan



Prof. Ir. Danny Manongga, M.Sc., Ph.D.
Ketua Program Studi MSI

Dinyatakan Lulus Tanggal : 15 Oktober 2018

Penguji :

- Dr. Ade Iriani, M.M



- Hindriyanto Dwi Purnomo, ST., MIT., Ph.D.



Kata Pengantar

Puji syukur Penulis panjatkan kehadiran Tuhan YME. Atas rahmat dan hidayahnya sehingga dapat menyelesaikan penelitian yang berjudul “**Analisis Data Status Di Facebook Untuk Merekomendasikan Produk Makanan Ringan Menggunakan *Tf-Idf* Dan *Document Weighting***”, yang diajukan sebagai persyaratan dalam menyelesaikan Program Studi S2 di Magister Sistem Informasi pada Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana dan meraih gelar Magister Komputer.

Dalam kesempatan ini Penulis mengucapkan terima kasih sebesar-besarnya kepada semua pihak yang telah membantu, hingga terselesaikannya penulisan laporan penelitian ini. Ucapan terima kasih penulis berikan kepada :

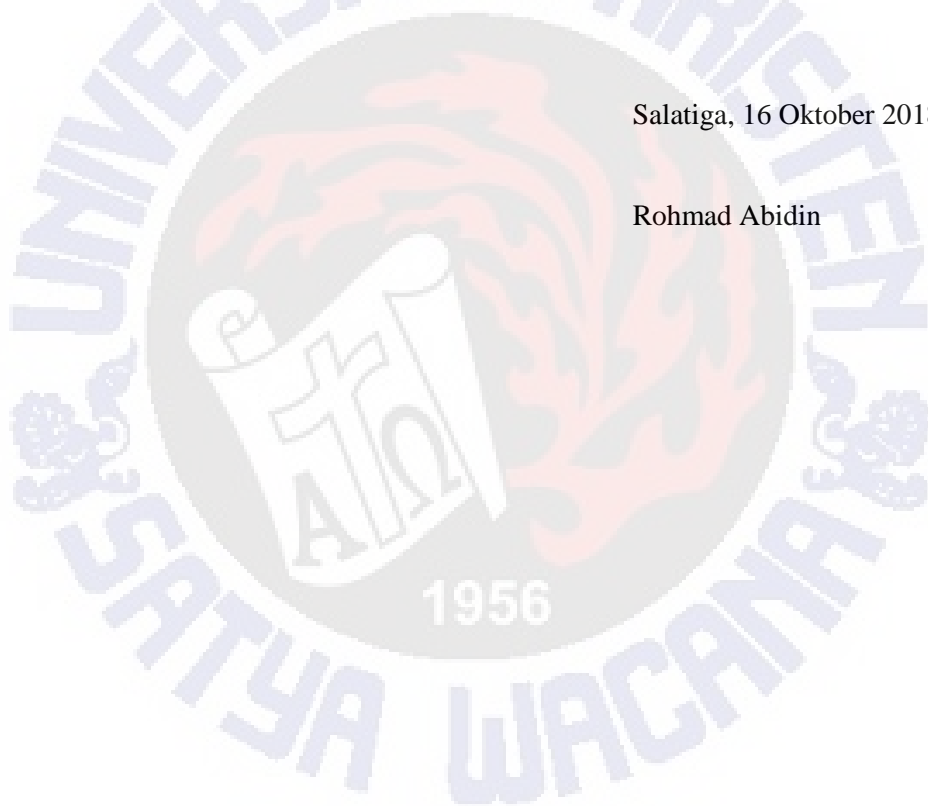
1. Bapak Wiwin Sulisty, ST., M.Kom., selaku Dekan Fakultas Teknologi Informasi Universitas Kristen Satya Wacana Salatiga.
2. Bapak Prof. Ir. Danny Manongga, MSc., Ph.D., selaku Ketua Program Studi Magister Sistem Informasi, Fakultas Teknologi Informasi Universitas Kristen Satya Wacana Salatiga.
3. Bapak Prof. Dr. Ir. Eko Sedyono, M.Kom selaku pembimbing yang selalu ada dan membimbing penulis serta selalu ada waktu, tenaga dan juga saran-saran dalam penulisan.
4. Bapak Dr. Sri Yulianto Joko Prasetyo, S.Si., M.Kom selaku pembimbing yang selalu memberikan motivasi serta jalan keluar bagi kesulitan-kesulitan yang penulis hadapi.
5. Orang tua tercinta, kakak dan adik yang selalu memberikan do’a dan semangat penulis.
6. Istriku Nila Auliyani, AH., serta anak-anakku tersayang yang selalu berdo’a dan memberikan energi selama menyelesaikan studi.

7. Seluruh Bapak/Ibu dosen pengajar dan seluruh staff di Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana yang telah mendidik dan memberikan motivasi dalam penelitian ini.
8. Semua pihak yang telah membantu yang tidak dapat disebutkan satu demi satu.

Sangat disadari bahwa laporan penelitian ini masih jauh dari kesempurnaan. Oleh sebab itu, kritik dan saran sangat diharapkan untuk melengkapi laporan penelitian ini, sehingga menjadi suatu bahan pembelajaran yang dapat memajukan kreatifitas anak bangsa. Terima kasih semuanya.

Salatiga, 16 Oktober 2018

Rohmad Abidin



ABSTRAK

Tujuan dari penelitian ini adalah mengembangkan sistem analisis status pada akun media sosial untuk mendukung kegiatan promosi Usaha Mikro Kecil khususnya produk makanan ringan. Penelitian menggunakan teknik yang terdapat pada *text mining*, analisis *Term Frequency- Invers Document Frequency* dan *Document Weighting* untuk merekomendasikan produk. Objek penelitian akun media sosial facebook dengan *sample random* yang tergabung dalam grup Warga Demak serta produk hasil UMK yang terdaftar pada Dinas Perdagangan dan Koperasi Kabupaten Demak. Hasil akhir penelitian ini adalah *prototype* sistem analisis status media sosial untuk merekomendasikan produk UMK yang sesuai dengan hasil analisa status facebook. Dari hasil analisa didapatkan rata-rata tingkat akurasi prediksi sebesar 73.3% untuk keberhasilan merekomendasikan produk. Hasil penelitian ini menunjukkan bahwa analisis *TF-IDF* dan *Document Weighting* dapat dimanfaatkan untuk pengembangan sistem baru guna mendukung kegiatan promosi pengelola kelompok UMK agar lebih tepat sasaran.

Kata kunci: media sosial, *TF-IDF* , *document weighting*, UMK.

ABSTRACT

The purposes of the research are developing analysis application of status on social media account to support Micro, Small Enterprises (MSE) promotion especially on snack product. The research used techniques in text mining, Term Frequency-Invers Document Frequency analysis and Document Weighting to recommendation the product. The object of research is Facebook status updates as well as products of MSE who registered at the Department of Trade and Cooperatives in Demak District. The final result of this research is prototype social media status analysis system to recommend products from results of system analysis. From the analysis results obtained an average predictive accuracy rate of 73.3% for recommend the product. The results of this study it can be concluded that the analysis of TF - IDF and Document Weighting can be utilized for the development of new systems to support the promotion activities of MSE group managers to be more effective.

Keywords: social media, *TF-IDF*, *document weighting*, MSE

1. PENDAHULUAN

Setiap orang memiliki motif berbeda-beda dalam menggunakan Media sosial. Utomo melakukan penelitian mengenai motif penggunaan media sosial mendapatkan bahwa motif tertinggi orang menggunakan media sosial adalah untuk memperbaharui status, dimana mereka mampu menyatakan eksistensi dirinya serta mendapatkan informasi[1].

Media sosial dapat digunakan untuk menunjang aktivitas keseharian pengguna. Seperti perusahaan atau individu menggunakan media sosial untuk menunjang aktivitas bisnis, saling berbagi karya dan sebagai tempat *sharing* untuk yang memiliki kesamaan hobi. Dengan adanya media sosial di aplikasi *mobile* membuat setiap orang bisa menggunakan media sosial dimana saja sepanjang terkoneksi jaringan internet [2].

Media sosial dapat dijadikan sebagai *Integrated Marketing Communication* (IMC) guna memberikan keunggulan kompetitif bagi UKM. Setidaknya UKM dapat melakukan komunikasi kepada pelanggan yang telah ada dan menarik calon pelanggan baru, agar UKM dapat bertahan ditengah persaingan [3].

Penelitian mengenai analisis media sosial pernah dilakukan Muhammad dalam penelitiannya berjudul *Penerapan Data Mining Dalam Menganalisis Kepribadian Pengguna Media Sosial Dengan Naïvebayes Classifier*. Berdasarkan teori *five factor model*, hasil penelitian menyimpulkan bahwa kepribadian *Openness to New Experiences* sebanyak 11.68%, Kepribadian *Conscientiousness* 5.34%, kepribadian *Extraversion* 5.5%, kepribadian *Agreeableness* 8.54%, kepribadian *Neuroticism* 4.82%, serta 64.12% tidak dapat diklasifikasikan [4].

Untuk penelitian efektifitas iklan pada jejaring sosial yang dilakukan oleh Hasanah mengemukakan hasil dari perhitungan EPIC *rate* yaitu iklan Facebook Huma Harati dinyatakan efektif sebagai media promosi bagi Huma Harati. Namun pada skala efektif menunjukkan bahwa promosi yang dilakukan Huma Harati Kalteng melalui iklan di Facebook belum mencapai hasil yang maksimal, oleh karena itu Huma Harati Kalteng perlu meningkatkan promosi yang lebih baik lagi [5].

Penelitian lain tentang promosi menggunakan media sosial dilakukan oleh Priambada yang menghasilkan bahwa media sosial yang paling banyak digunakan adalah web blog, facebook dan E-mail. Manfaat yang paling dirasakan adalah media sosial merupakan sarana komunikasi yang efektif, dapat meningkatkan pangsa pasar serta membantu keputusan bisnis. Manfaat penggunaan media sosial untuk mengadakan kontak personal dengan konsumen, promosi/advertising, mendata kebutuhan konsumen dan menyampaikan respon ke konsumen [6].

Untuk mengenal calon konsumen secara lebih personal, penelitian ini berusaha menggabungkan data hasil text mining status di media sosial seseorang yang di integrasikan dengan data UKM. Penelitian ini akan menghasilkan sebuah *prototype* sistem rekomendasi produk UMK sesuai dengan hasil analisa status pada *timeline* media sosial masing-masing akun.

2. TINJAUAN PUSTAKA

2.1 Kajian Literatur

Penelitian penggunaan *TF-IDF* dalam mengklasifikasikan pengguna media sosial pernah dilakukan oleh Rivki dengan menggunakan *text mining*, analisis *TF-IDF* serta Algoritma *K-Nearest Neighbor* peneliti mengklasifikasi ketertarikan follower berdasarkan status di Twitter. Hasil nilai akurasi terbesar yang didapat pada proses klasifikasi untuk empat kali pengujian adalah 68% [7].

Ardi melakukan penelitian dengan membandingkan performa *TF-IDF* dengan n-gam menggunakan *Support Vector Machine* (SVM) untuk mendapatkan hasil klasifikasi terbaik untuk menganalisis sentiment masyarakat terhadap *marketplace*. Peneliti menemukan bahwa Unigram memiliki nilai akurasi tertinggi, dengan nilai akurasi 80,87% [8].

Chang menggunakan metode *TF-IDF* dan *Social Network Analysis* dalam sebuah studi kasus menyajikan bagaimana sistem rekomendasi bekerja berdasarkan data yang dikumpulkan dari Plurk. Dari hasil analisis, peneliti menemukan perbedaan pengukuran SNA antara berbagai produk yang signifikan, hal itu menunjukkan bahwa sistem rekomendasi dapat dikerjakan untuk merekomendasikan produk yang berbeda untuk menargetkan pelanggan [9].

2.2 Text Mining

Text mining merupakan salah satu penerapan konsep dan teknik pada data mining untuk mencari pola di dalam teks, atau proses menganalisis teks guna mengekstrak informasi yang bermanfaat untuk tujuan tertentu. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Sumber data yang digunakan pada *text mining* adalah kumpulan teks yang memiliki format yang tidak terstruktur atau setidaknya semi terstruktur. Penggunaan *text mining* antara lain dalam pengkategorisasian teks (*text categorization*) dan pengelompokan teks (*text clustering*). Terdapat beberapa tahapan dalam melakukan text mining, yaitu :

a. Case Folding

Pada tahap ini semua huruf dalam dokumen diubah menjadi *lower case* (huruf kecil). Hanya huruf 'a' sampai dengan 'z' yang diterima. Karakter selain huruf dihilangkan dan dianggap *delimiter*.

b. Tokenizing

Tahap *tokenizing / parsing* adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya [10]. Teknik token menggunakan fungsi *split* dimana pemisahan string dilakukan berdasarkan *white space* (spasi dan tab). Tokenizing berfungsi memecah suatu kalimat menjadi sekumpulan kata untuk memudahkan proses *filtering* dan *stemming*.

c. Filtering

Filtering dilakukan dengan metode *Stopword Removal* yaitu proses yang dilakukan untuk menghilangkan kata-kata yang sering muncul dan tidak penting untuk di tampilkan. Contoh: saya, aku, kamu, dan, bagaimana, !, '? dan lain sebagainya.

d. Stemming

Tahap *Stemming* merupakan proses mencari akar kata dari tiap kata. Proses *stemming* dokumen teks menggunakan kamus yang sudah disediakan, jika tidak ada maka akan di proses dengan Algoritma Nazief & Adriani. Pemilihan Algoritma ini berdasarkan hasil penelitian Wahyudi bahwa algoritma Nazief & Adriani memberikan hasil lebih baik untuk stemming pada dokumen berbahasa Indonesia dengan akurasi 95,26% dibandingkan algoritma Porter dengan akurasi 79,13% [11].

2.3 Term Frequency (TF)

Term frequency merupakan salah satu metode yang digunakan untuk melakukan perhitungan pembobotan term. Fitur *term frequency* dilakukan dengan menghitung frekuensi kemunculan kata (*term*) tertentu pada suatu dokumen. Semakin besar frekuensi kemunculan suatu *term* di dalam suatu dokumen, akan memberikan nilai kesesuaian yang semakin besar. *Term frequency* merupakan metode *feature selection* yang sederhana dengan waktu komputasi yang rendah tetapi memiliki hasil yang akurat. Algoritma Term Frequency yang dipakai adalah *Raw TF*. *Raw TF*, memberikan nilai TF berdasarkan jumlah kemunculan suatu *term* di dokumen. Sebagai contoh jika *term* muncul 3 kali maka kata tersebut akan bernilai 3. Koleksi dokumen yang besar *Raw Term Frequency* memiliki nilai kesesuaian lebih tinggi dibanding algoritma *Term Frequency* yang lain. [12]

2.4 Invers Document Frequency (IDF)

Suatu *term* unik yang menunjukkan ke khas-an suatu produk kemungkinan akan sedikit muncul pada status facebook seseorang. Sehingga, guna mengurangi bobot kata-kata yang

bersifat umum dan meningkatkan bobot *term* unik maka digunakan *Inverse Document Frequency* (IDF). Rumus menghitung IDF sebagai berikut :

$$idf_t = \log \left(\frac{N}{df_t} \right)$$

Dimana :

N = total dokumen

df = banyak dokumen yang mengandung kata yang dicari

t = term

2.5 Document Weighting

Setelah didapatkan nilai IDF masing-masing *term*, untuk mengetahui produk yang sesuai maka diperlukan pembobotan masing-masing dokumen produk (*Document Weighting*). Menghitung bobot setiap dokumen (d) untuk setiap *term* (t) menggunakan rumus:

$$W_{dt} = tf_{dt} * IDF_t$$

Dimana :

d : dokumen ke- d

t : kata ke- t dari kata kunci

W : bobot dokumen ke- d terhadap kata ke- t

tf : banyaknya kata yang dicari pada sebuah dokumen

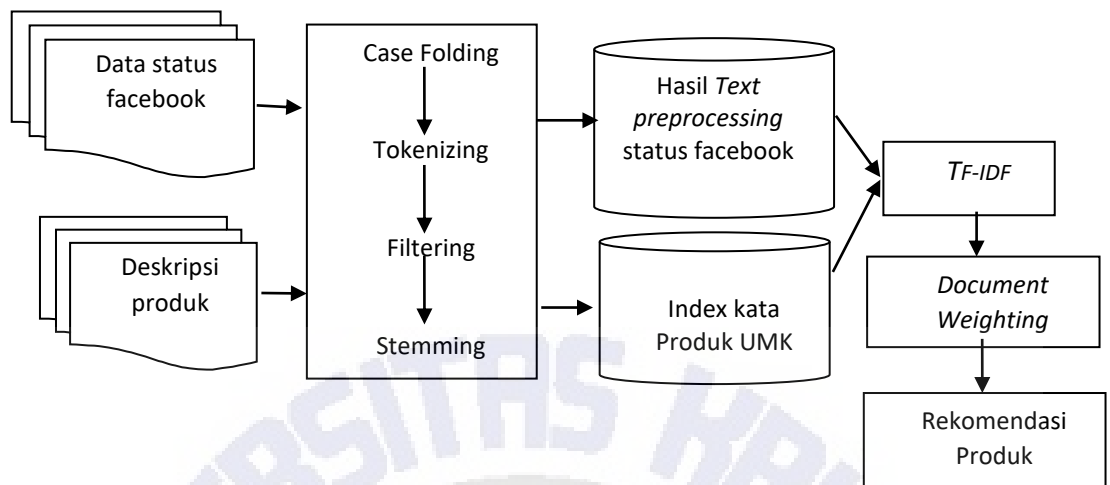
Setelah bobot masing-masing dokumen diketahui, maka dilakukan proses pengurutan dimana semakin besar nilai W , semakin besar tingkat similaritas dokumen tersebut terhadap kata yang dicari, demikian pula sebaliknya.

2.6 Metodologi Penelitian

Penelitian ini dilaksanakan dalam tiga tahap yaitu: a) pengumpulan data penelitian b) analisis data c) pengujian hasil analisis produk. Pemrosesan awal data penelitian mencakup aktivitas pengumpulan data produk UMK serta koleksi status akun media sosial. Untuk analisis data terbagi menjadi 2 proses penelitian yaitu *text mining processing*, dan *data comparison*. *Text mining processing* mengolah data input berupa status media sosial serta deskripsi produk UMK. Proses *text mining* dilakukan menggunakan metode yang secara umum dilakukan pada *text mining* yaitu tahap *tokenizing*, *case folding*, *filtering*, serta *stemming*. Proses *text mining* dilakukan pada deskripsi produk makanan ringan yang akan menghasilkan *index* kata pada setiap produk. Proses *text mining* data status media sosial dimulai dengan mengumpulkan semua status facebook calon reponden dari awal mendaftar facebook hingga terakhir, kemudian dipilih status yang mengandung kata-kata yang berhubungan makanan ringan seperti makan, lezat, enak, gurih, dan lain-lain.

Proses yang kedua adalah *data comparison*, yaitu membandingkan antara data hasil *text mining* akun media sosial dengan *index* kata pada produk UMK. Proses perbandingan menggunakan *TF-IDF*. Hasil dari *TF-IDF* digunakan untuk memberikan peringkat produk yang terdapat dalam koleksi produk menggunakan kaidah *Document Weighting*. Dari seluruh produk yang terdaftar diambil 5 produk tertinggi untuk di promosikan kepada calon konsumen.

Peneliti melakukan ujicoba pertama dengan menggunakan kuesioner untuk menentukan tingkat efektifitas penerapan *TF-IDF* dan *Document Weighting*. Berikut ini adalah diagram tahapan penelitian yang dilakukan (gambar 1).



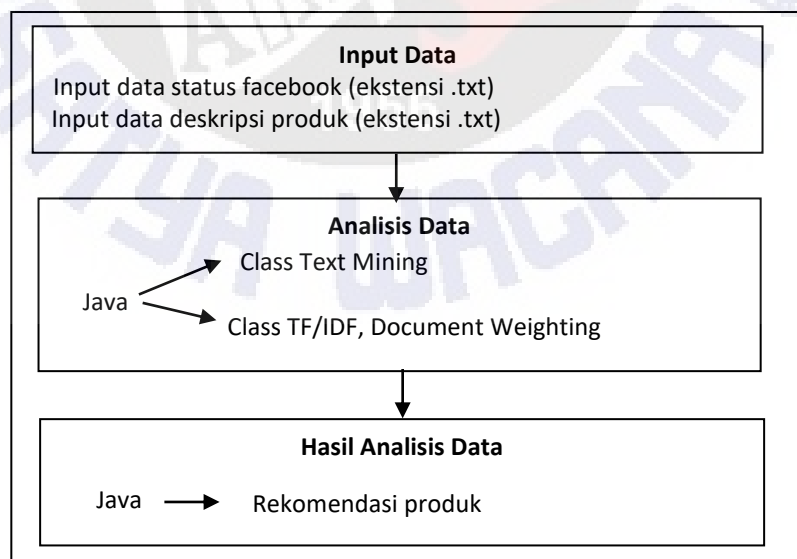
Gambar 1. Diagram Tahap Penelitian

3. HASIL DAN PEMBAHASAN

3.1 Hasil Penelitian

Desain arsitektural sistem analisis ini dapat dilihat pada Gambar 2. Desain arsitektural sistem terdapat 3 kelompok komponen pada aplikasi yaitu : a) *input data*, b) analisis data c) *output* rekomendasi produk.

Pada gambar 2 memperlihatkan bahwa program memerlukan input data status facebook yang di-copy secara manual dan data deskripsi produk UMK yang disimpan format .txt kemudian dianalisis oleh sistem dengan tahapan yang sudah ditentukan dalam *text mining*. Setelah proses *text mining* selesai, hasil data status facebook dibandingkan dengan data produk UMK menggunakan proses pencairan *TF / IDF* dan *Document Weighting*. Terakhir sistem akan menghasilkan teks rekomendasi produk yang dapat di unduh untuk disimpan.



Gambar 2. Desain Arsitektur

Tampilan dibuat sederhana agar pengguna dapat dengan mudah menggunakan sistem ini. Pada aplikasi juga terdapat halaman utama untuk menganalisis status facebook (Gambar 3). Proses analisis sistem dilakukan dengan tahapan yang telah ditentukan. Proses dilakukan pada *background* sehingga tidak terlihat oleh pengguna. Hasil analisis status facebook akan memberikan rekomendasi produk dengan *sorting* skor tertinggi. Terdapat juga tombol untuk pengaturan produk (Gambar 4) yang berfungsi untuk menambah atau mengurangi produk yang akan dianalisa.

Gambar 3. Form Analisis Data

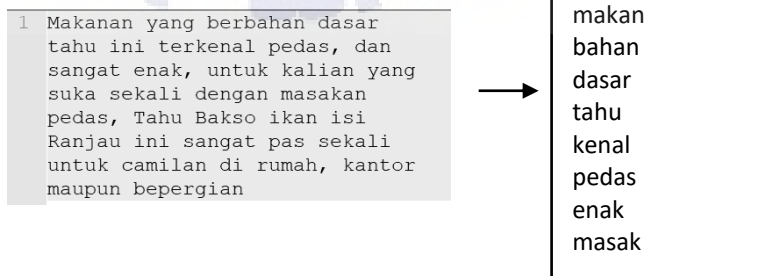
Gambar 4. Form Pengaturan Produk

3.2 Perhitungan *TF-IDF* dan *Document Weighting*

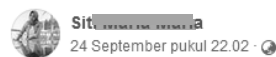
Terlebih dahulu dilakukan *text preprocessing* pada masing-masing deskripsi produk untuk menghasilkan *index* kata tiap produk kemudian dimasukkan ke dalam *database*. Setelah semua data produk tersimpan, selanjutnya data status facebook diproses dengan tahapan yang sama dengan data produk. Data yang status facebook yang sudah dilakukan proses *text preprocessing* diatas kemudian dilakukan perhitungan *TF – IDF* dan *document weighting*.

Langkah analisis dilakukan dengan tahapan sebagai berikut :

1. Langkah pertama dilakukan *text preprocessing* pada produk UMK untuk menghasilkan *index* kata pada masing-masing produk. *Text preprocessing* juga dilakukan pada data status facebook. Tujuan utama dari *text preprocessing* adalah menghilangkan kata-kata tidak penting serta mengubah kata menjadi kata dasar agar mudah digunakan untuk pencarian kata yang sama di masing-masing deskripsi produk. Di bawah ini menunjukkan proses *text preprocessing* pada deskripsi produk (Gambar 5) dan status facebook (Gambar 6).



Gambar 5. Hasil *text preprocessing* deskripsi produk



Nyam nyam nyam 😊 dimasak apa aja...aku suka 😊 @bakso mercon pedas, bakso pete taoco pedas, bakso sate bakar
Kalian sukanya yg manaaaa...

masak
bakso
mercon
pedas
pete
taoco
sate
bakar

Gambar 6. Hasil text preprocessing status facebook

2. Hitung jumlah kata (*term frequency*) dengan mencari kata yang sama antara data status facebook dengan masing-masing deskripsi produk (*D*). Hasil *term frequency* akan menunjukkan bahwa semakin banyak kata yang sama antara status facebook dengan deskripsi produk maka semakin besar kemungkinan pemilik akun memilih produk tersebut.

Pemberian bobot kata pada dokumen menggunakan kode sebagai berikut :

```
private Double tf(List<String> doc, String term) {
    double tf = 0;
    for (String word : doc) {
        if (term.equalsIgnoreCase(word)) {
            tf++;
        }
    }
    System.out.print(" Raw TF : " + tf);
    return tf;
}
```

3. Tentukan jumlah kata yang memiliki nilai di setiap produk (*dokumen frequency/df*). Jika di dalam produk memiliki nilai *tf* berapapun nilainya maka akan diberi nilai 1, jika tidak memiliki nilai *tf* maka diberi nilai 0.

$$\begin{aligned}
 df_{i3} &= d1_{i3} + d2_{i3} + d3_{i3} + \dots + d15_{i3} \\
 &= 1 + 0 + 1 + 1 + 1 + 1 + 0 + 0 + 0 + 0 + 0 + 0 + 1 + 1 + 1 \\
 &= 10
 \end{aligned}$$

Pencarian nilai *df* bertujuan untuk mengetahui sebaran suatu kata, semakin banyak tersebar di deskripsi produk yang lain maka kata tersebut dianggap kata yang umum, semakin sedikit sebaran suatu kata di produk lain, maka kata tersebut dapat menunjukkan kekhasan suatu produk.

4. Hitung nilai *IDF* dari tiap-tiap *term*, fungsi rumus menghitung *IDF* sebagai berikut :

$$\begin{aligned}
 idf_{i3} &= \text{LOG}\left(\frac{d}{df_{t,13}}\right) \\
 &= \text{LOG}\left(\frac{15}{10}\right) \\
 &= \text{LOG}(1,5) \\
 &= 0.176091259
 \end{aligned}$$

Fungsi *IDF* diatas guna mengurangi bobot kata-kata yang bersifat umum dan meningkatkan bobot kata unik. Semakin tinggi bobot kata, di suatu produk semakin besar kemungkinan produk akan dipilih. Berikut *listing* program menghitung *Document Frequency* dan *Invers Document Frequency*.

```
private double idf(List<List<String>> docs, String term) {
    double f = 0;
    for (List<String> doc : docs) {
        for (String word : doc) {
            if (term.equalsIgnoreCase(word)) {
                f++;
                break; }
        }
    }
    System.out.print(" DF : "+f);
    if(f==0){
        System.out.print(" IDF : 0");
        return 0;
    }else{
        System.out.print(" IDF : "+Math.log10(docs.size() / f));
        return Math.log10(docs.size() / f);
    }
}
```

5. Lakukan pembobotan tiap kata pada deskripsi produk (*d*) yang telah disimpan. Pemberian bobot kata pada dokumen menggunakan rumus sebagai berikut :

$$\begin{aligned}
 W_{di3} &= tf_{i3} \times idf_{i3} \\
 &= 1 \times 0.176091259 \\
 &= 0.176091259
 \end{aligned}$$

Berikut *listing* kode untuk menghitung bobot tiap kata.

```
private double Wdt(List<String> doc, List<List<String>> docs, String term)
{
    return tf(doc, term) * idf(docs, term);
}
```

6. Jumlahkan tiap Bobot W_{di} untuk mengetahui nilai bobot tiap produk (W_d) semakin besar bobot dokumen, kemungkinan semakin tinggi dipilih oleh konsumen.

$$\begin{aligned}
 W_{d1} &= W_{d11} + W_{d12} + W_{d13} + \dots + W_{d125} \\
 &= 0 + 0 + 0.176091259 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0.574031268 + 0 + 0 + 0 + \\
 &\quad 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \\
 &= 0.750122527
 \end{aligned}$$

Berikut *listing* kode menghitung bobot setiap produk (W_d).

```
Double[][] hasilWeight = new Double[documents.size()][documents.size()];
TFIDFWeight calculator = new TFIDFWeight();
Double[] docWeight = new Double[documents.size()];
for (int a = 0; a < documents.size(); a++) {
    System.out.println("dokumen ke "+(a+1));
    docWeight[a] = 0.0;
    System.out.println("dokumen ke "+(a+1));
    for (int i = 0; i < medsos.size(); i++) {
        System.out.print(medsos.get(i));
        double Wdt = calculator.Wdt(documents.get(a), documents,
            medsos.get(i));
        //tambahkan bobot produk tiap ada term
        docWeight[a] = docWeight[a] + Wdt;
    }
}
```

```

    }
    hasilWeight[a][0] = Double.valueOf(a);
    hasilWeight[a][1] = docWeight[a];
    System.out.println(" Dokumen Weighting = " + docWeight[a]);
}

```

Baris kode pada program diatas menunjukkan perhitungan bobot tiap kata terdapat dokumen akan dijumlahkan hingga selesai untuk satu produk. Jika produk pertama selesai dianalisis akan dilanjutkan ke produk kedua dan seterusnya hingga semua produk selesai di analisis.

Hasil perhitungan pada bobot produk pertama (W_{d1}) memiliki bobot 0.750122527. Hasil perhitungan bobot dari setiap produk disajikan pada tabel 1 di bawah ini :

Tabel 1. Hasil Pemeringkatan Bobot Produk

Kode	Nama Produk	Bobot	Ranking
D2	bintang 5	7.1077	1
D14	rangin super	2.801275	2
D9	keripik pisang raja	1.528274	3
D8	keripik pisang enggano	1.449093	4
D12	kerupuk ikan laut	1.176091	5

Dari hasil pemeringkatan diatas maka dipilih 5 produk tertinggi untuk direkomendasikan kepada pelanggan untuk ditawarkan.

3.3 Analisa Hasil

Peneliti melakukan uji coba dengan menggunakan kuesioner untuk menentukan tingkat efektifitas sistem yang dihasilkan. Populasi objek penelitian adalah anggota grup facebook Warga Demak yang telah mempunyai akun facebook lebih dari 5 tahun dan berusia lebih dari 17 tahun. Sedangkan jumlah produk yang digunakan untuk bahan ujicoba dibagi menjadi 3 tahap pengujian yaitu tahap 1 sebanyak 15 produk, tahap 2 sebanyak 30 produk, serta tahap 3 sebanyak 50 produk. Masing-masing tahap sebanyak 50 responden. Pada tabel 2 menunjukkan hasil uji coba sistem yang dihasilkan :

Tabel 2. Hasil Analisa

Jumlah Produk	Memilih	Tidak memilih
15	67%	30.3%
30	75.6%	24.4%
45	77.2%	22.8%
Rata-rata	73.3%	26.7%

4. KESIMPULAN

4.1 Kesimpulan

Dari hasil pertama menunjukkan bahwa tingkat akurasi yang dihasilkan oleh algoritma *TF-IDF* dan *Document Weighting* dalam merekomendasikan produk sebesar 73.3% dengan nilai *recall* 93%. Semakin banyak produk yang digunakan dalam penelitian semakin besar nilai *precision*. Sehingga dapat disimpulkan bahwa metode *TF-IDF*, *Document Weighting* dapat digunakan sebagai algoritma untuk merekomendasikan produk pada pelanggan sesuai dengan yang diminati melalui analisis status akun facebook.

4.2 Saran

Saran dalam pengembangan sistem kedepan sebagai berikut :

1. Penambahan berbagai jenis produk yang lebih beragam tidak hanya makanan ringan.
2. Kedepan perlu dikembangkan pula meneliti jaringan grup yang diikuti agar lebih maksimal dalam merekomendasikan produk.
3. Penambahan metode baru untuk mengelompokkan pengguna berdasarkan keseragaman produk.

DAFTAR PUSTAKA

- [1] Utomo, Dea Anggraeni. 2013. *Motif Pengguna Jejaring Sosial Google+ Di Indonesia*, Jurnal E-Komunikasi Vol I. No.3 Tahun 2013
- [2] Luik, Jandy E. 2011. *Media Sosial dan Presentasi Diri, Bab Buku Komunikasi 2.0: Teoritisasi dan Implikasi*. Yogyakarta: AspiKom
- [3] Siswanto, T., 2013. Optimalisasi sosial media sebagai media pemasaran usaha kecil menengah. *Jurnal Liquidity*, 2(1), pp.80-86.
- [4] Muhammad, Harits. 2016. *Penerapan Data Mining dalam Menganalisis Kepribadian Pengguna Media Sosial dengan Naive Bayes Classifier: Studi Kasus Media Sosial Instagram*. Jurnal Informatika Universitas Padjadjaran. Vol. 01 No. 01 Agustus 2016. ISSN : 2503-5258
- [5] Hasanah, Nur, dkk., 2016. *Analisis Efektivitas Iklan Jejaring Sosial sebagai Media Promosi Menggunakan EPIC Model*. Scientific Journal of Informatics, [S.l.], v. 2, n. 2, p. 99-109, feb. 2016. ISSN 2460-0040.
- [6] Priambada, S., 2015. *Manfaat Penggunaan Media sosial Pada Usaha Kecil Menengah (UKM)*. SESINDO 2015, 2015.
- [7] Rivki, M., & Bachtiar, A. M., 2017. *Implementasi Algoritma K-Nearest Neighbor Dalam Pengklasifikasian Follower Twitter Yang Menggunakan Bahasa Indonesia*. *Journal of Information System*, 13(1).
- [8] Ardi, H. L., Sedyono, E., & Ningrum, R. K., 2017. Support Vector Machine Classifier For Sentiment Analysis Of Feedback Marketplace For Sentiment Analysis Of Feedback Marketplace With A Comparison Features At Aspect Level. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, 4(11), 7–12. <https://doi.org/10.26562/IJIRAE.2017.NVAE10080>
- [9] Chang, P., Ting, I.-H., & Wang, S., 2011. Towards Social Recommendation System Based on the Data from Microblogs. In *International Conference on Advances in Social Networks Analysis and Mining* (pp. 672–677). IEEE. <https://doi.org/10.1109/ASONAM.2011.101>
- [10] Triawati, Chandra. 2009. *Metode Pembobotan Statistical Concept Based untuk Klastering dan Kategorisasi Dokumen Berbahasa Indonesia*, Institut Teknologi Telkom Bandung. Teknologi Telkom Bandung.
- [11] Wahyudi, Dwi. 2017. *Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani dan Porter Pada Dokumen Berbahasa Indonesia*. Jurnal Sinus Vol 15, No 2. ISSN (Online) : 2548-4028
- [12] Wibowo Ari. 2011. *Pengujian Kerelevanan Sistem Temu Kembali Informasi*. Seminar Nasional Ilmu Komputer, Universitas Gadjah Mada, 19 November 2011.